Progress Report

Period Covered by the Report
• *11 December 2001 to 8 January 2002*

Date of Report
• *9 January 2002*

Project
• ***Compadre: A Device Independent Voice-to-Voice Language Translator Software Solution***
• SBIR Phase I, Topic N01-044
• Contract Number N00014-01-M-0225, Amendment P00001

Item Number
0002AA: Progress Report

Security Classification
Unclassified

Sponsor
Dr. Joel Davis
Office of Naval Research
Ballston Tower One
800 North Quincy Street
Arlington, VA 22217-5660

Principle Investigator:
Robert Palmquist
SpeechGear, Inc
3454 Circle Bluff Court
Faribault, MN 55021

Program Partners

VISUAL GOLD

IRISUSA

ARAMEDIA

L&H/Apptek

MSU

A. Project Summary

Technical Abstract:

Mission Statement
To develop and deploy language translation software that is device independent, supports bi-directional translation of multiple languages, produces text transcriptions of spoken conversations and supports translation of text extracted from digital images. This software shall run in both a reduced functionality standalone mode, and by wirelessly connecting to remote servers, a full-function mode. This software shall run on multiple pocketable platforms resulting in a mobile system that is low in cost, easy to use, robust in operation and comfortable to carry and/or wear.

The object of this Phase I research effort is to investigate the scientific, technical and commercial merit and feasibility of the system described in the preceding mission statement. Specifically, the team will investigate design options for the mobile translator system, identify potential applications, and select the best option(s) to pursue in making the design a reality. Four technical areas will be investigated: potential pocketable computing platforms, the operator interface, optical character recognition software and the language translation software. The commercial feasibility of this design will also be investigated. This includes identifying potential applications, languages to be supported, cost, and user requirements such as interface modes and response times. By combining both the commercial and technical elements, a complete definition of successful software and system solutions for pocketable language translation devices will be achieved.

Prototype systems showing device independence will be developed and demonstrated and a final report written documenting the Phase I results and recommendations for follow-on research and development in Phase II. Options are included for incorporating additional language pairs into the system and application specific terminology.

Anticipated Benefits/Potential Commercial Applications of the Research or Development:

Applications include all individuals who require multi-lingual capabilities. The mobile translator will benefit a wide range of individuals including military personnel, airport employees, border patrol and customs agents, police, fire fighters, retail clerks, bank tellers, delivery personnel, phone operators, tourists and any industry that sells, develops or manufactures products to/in global markets or employs individuals that do not speak the native language.

A.4 Camera-Based Mode

A4.1 -- Brief Summary

The primary means to input text into the SmartPhone for this mode of usage will be a digital camera. Such a system is shown in Figure CM1. The digital camera will be used to capture an image of the foreign language. Such a picture is shown in Figure CM2.

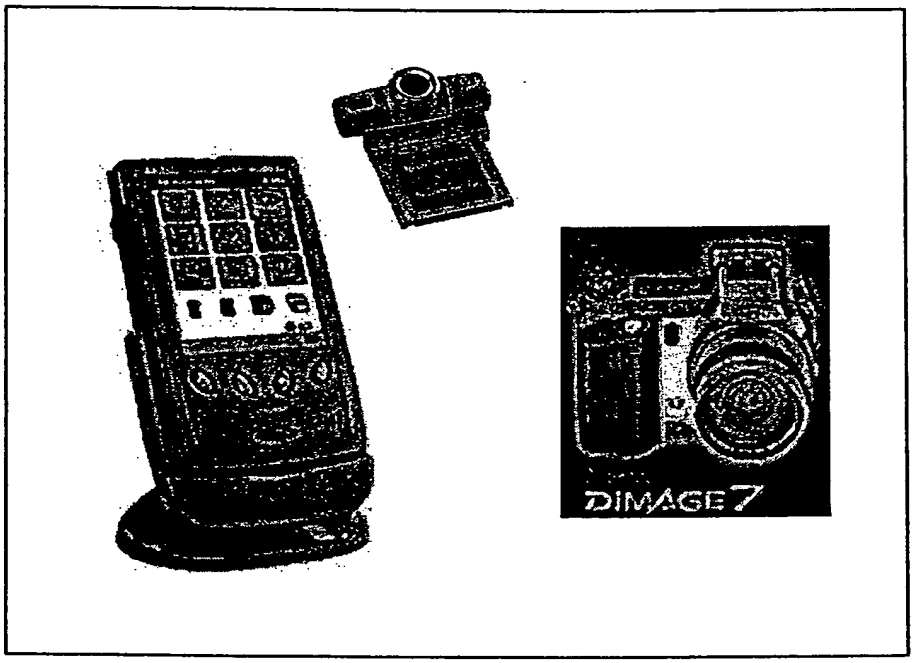

Figure CM1: Examples of Camera-based Systems

Once the desired image is obtained, the SmartPhone will wirelessly connect to a remote server where the image will be processed and the resulting translation sent back to the user. An example of the translated text in the proposed "one-click" GUI is shown in Figure CM3. For most applications, this connection will be made using cellular telephones. Because of the limited bandwidth of such a connection, it is important to reduce the overall size of the transmission. Thus, SpeechGear is in the process of evaluating different image compression algorithms. These algorithms will be embedded directly into SpeechGear's software, and thus will be transparent to the end user.

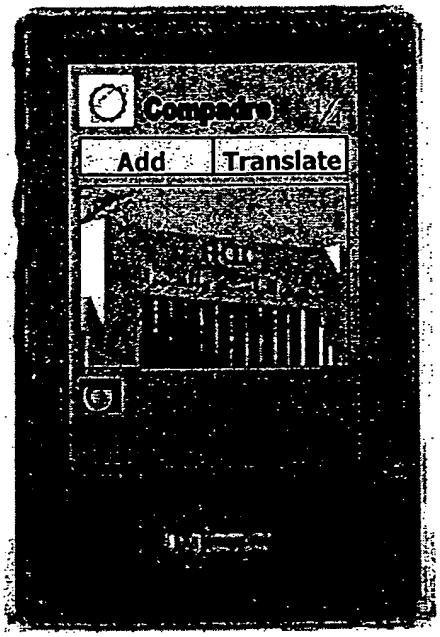

Figure CM2: Example Graphical User Interface to Submit Images for Translation

As is shown in Figure CM2, a "one-click" GUI is planned. After capturing the image(s), the user will simply select "Translate" and the wireless connection will automatically be established. Note that multiple images can be sent simultaneously using a single click. This is similar to the "Add to Basket" interfaces that are being used at web-based shopping sites. In this approach, items that are selected can be loaded into a virtual basket or cart, and once you are done shopping you can select "Check Out" to purchase all of the items simultaneously. For *Compadre*, multiple images can be selected and entered into the queue, and when the user is ready to connect to the remote server, then simply selecting the "Translate" button will connect the SmartPhone to the remote server, which in turn will process the images and return the resulting translation. The images will be transmitted back to the user using an HTML format. The users can then scroll through these images and save or delete them as is desired. Please note that the actual buttons will be Icons versus text, and thus the look and feel of the resulting GUI will be a substantial improvement over what is shown in the Figures.

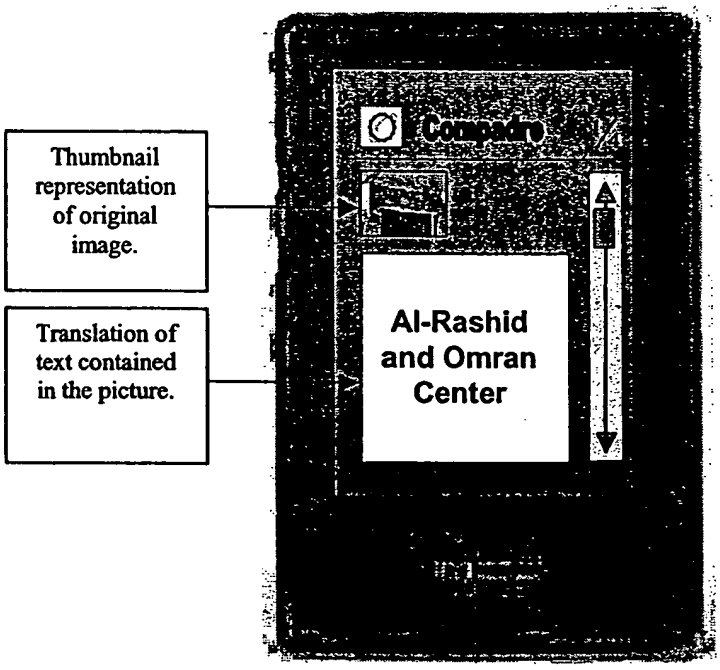

Figure CM3: Example of Graphical User Interface for Viewing Results of Translation

A4.2 – System Requirements:

CM1. System shall support connectivity to a digital camera (1.0a).

CM2. System shall be capable of displaying the captured digital image (1.0a).

CM3. System shall allow the user to select region of text to be translated (1.0c).

CM4. System shall allow the user to use "One-Click" to transfer image to remote server for processing (1.0b).

CM5. System shall allow the user to add multiple images to the send buffer (1.0c).

CM6. System server shall be capable of processing multiple images on a single connection/transmission (1.0c).

CM7. System shall support wireless connectivity such as a cellular telephone (1.0a).

CM8. System shall include image compression algorithms to reduce transmission connect time (1.0c).

CM9. User shall have the ability to turn on/off the image compression capability (1.0c).

CM10. The returned image shall include a "thumbnail" picture of the original image along with the translated text (1.0a).

CM11. The user shall be capable of saving this return image on the SmartPhone (1.0b).

CM12. The user shall be capable of scrolling through multiple return images using a "one-click" interface (1.0b).

-- Note: See Table CM1 for a summary of requirements CM13 through CM44.

CM13. The system shall support bi-directional translation for English/Arabic (1.0a).

CM14. The system shall support bi-directional translation for English/Korean (1.0a).

CM15. The system shall support bi-directional translation for English/Japanese (2.0a)

CM16. The system shall support bi-directional translation for English/Spanish (2.0b)

CM17. N/A

CM18. The system shall support bi-directional translation for English/Serbian (2.0c)

CM19. The system shall support bi-directional translation for English/Mandarin Chinese (2.0c)

CM20. The system shall support single-directional translation for Mandarin Chinese to English (2.0b)

CM21. The system shall support single-directional translation for Serbian to English (2.0b)

CM22. The system shall support bi-directional translation for English/Albanian (3.0b)

CM23. The system shall support single-directional translation for Albanian to English (2.0b)

CM24. The system shall support bi-directional translation for English/Thai (2.0c)

CM25. The system shall support single-directional translation for Thai to English (2.0b)

CM26. The system shall support bi-directional translation for English/Creole (3.0b)

CM27. The system shall support singledirectional translation for Creole to English (2.0b)

CM28. The system shall support bi-directional translation for English/ Indonesian (2.0c)

CM29. The system shall support single-directional translation for Indonesian to English (2.0b)

CM30. The system shall support bi-directional translation for English/ Turkish (2.0c)

CM31. The system shall support single-directional translation for Turkish to English (2.0b)

CM32. The system shall support bi-directional translation for English/Malay (2.0c)

CM33. The system shall support single-directional translation for Malay to English (2.0b)

CM34. The system shall support bi-directional translation for English/Greek (2.0c)

CM35. The system shall support single-directional translation for Greek to English (2.0b)

CM36. The system shall support bi-directional translation for English/Russian (2.0c)

CM37. The system shall support single-directional translation for Russian to English (2.0b)

CM38. The system shall support bi-directional translation for English/French (2.0a)

CM39. The system shall support bi-directional translation for English/German (2.0a)

CM40. The system shall support bi-directional translation for English/Portuguese (2.0b)

CM41. The system shall support bi-directional translation for English/Hindustani (2.0c)

CM42. The system shall support single-directional translation for Hindustani to English (2.0b)

CM43. The system shall support bi-directional translation for English/Swedish (2.0b)

CM44. The system shall support bi-directional translation for English/Norwegian (2.0c)

Table CM1: Summary of Language Support Schedule

Language	Single Directional	Bi-Directional
Arabic	CM13 – 1.0a	CM13 – 1.0a
Korean	CM14 – 1.0a	CM14 – 1.0a
Japanese	CM15 – 2.0a	CM15 – 2.0a
Spanish	CM17 – 2.0b	CM16 – 2.0b
Serbian	CM18 – 2.0c	CM18 – 2.0c
Mandarin Chinese	CM20 – 2.0b	CM19 – 2.0c
Albanian	CM23 – 2.0b	CM22 – 3.0b
Thai	CM25 – 2.0b	CM24 – 2.0c
Creole	CM27 – 2.0b	CM26 – 3.0b
Indonesian	CM29 – 2.0b	CM28 – 3.0b
Turkish	CM31 – 2.0b	CM30 – 3.0b
Malay	CM33 – 2.0b	CM32 – 3.0b
Greek	CM35 – 2.0b	CM34 – 3.0b
Russian	CM37 – 2.0b	CM36 – 2.0c
French	CM38 – 2.0a	CM38 – 2.0a
German	CM39 – 2.0a	CM39 – 2.0a
Portuguese	CM40 – 2.0b	CM40 – 2.0b
Hindustani	CM42 – 2.0b	CM41 – 2.0c
Swedish	CM43 – 2.0b	CM43 – 2.0b
Norwegian	CM44 – 2.0c	CM44 – 2.0c

Note: If a discrepancy is present between the table entries and the line items, the line items take precedence.